

5	969	86.9	206	1	SOSD1_CHICK	Q6vya3	gallus gall
6	817.5	73.3	213	2	Q6VYA2_XENLA	Q6vya2	xenopus lae
7	811.5	72.8	213	2	Q688C5_XENTR	Q688c5	xenopus tro
8	796.5	71.4	213	2	Q6X4U2_XENLA	Q6x4u2	xenopus lae
9	741.5	66.5	187	2	Q66KQ4_XENLA	Q66kq4	xenopus lae
10	645	57.8	207	2	Q568H6_BRARE	Q568h6	brachydanio
11	615.5	55.2	153	2	Q6VYA1_XENLA	Q6vya1	xenopus lae
12	372.5	33.4	179	2	Q4RYK3_TETNG	Q4ryk3	tetraodon n
13	349.5	31.3	213	1	SOST_HUMAN	Q9bqb4	homo sapien
14	349.5	31.3	213	2	Q495N9_HUMAN	Q495n9	homo sapien
15	348	31.2	213	1	SOST_CERAE	Q9bg78	cercopithec
16	340	30.5	213	1	SOST_RAT	Q99p67	rattus norv
17	333	29.9	176	1	SOST_BOVIN	Q9bg79	bos taurus
18	331	29.7	211	1	SOST_MOUSE	Q99p68	mus musculu
19	115.5	10.4	361	2	Q4RJ69_TETNG	Q4rj69	tetraodon n
20	114	10.2	400	2	Q6PFL2_BRARE	Q6pfl2	brachydanio
21	113	10.1	385	2	Q6IR79_XENLA	Q6ir79	xenopus lae
22	111	10.0	369	2	Q6NWA0_BRARE	Q6nwa0	brachydanio
23	106.5	9.6	1034	2	Q35888_RAT	Q35888	rattus norv
24	101	9.1	461	2	Q658Q1_HUMAN	Q658q1	homo sapien
25	101	9.1	1233	1	MUC5A_HUMAN	P98088	homo sapien
26	100	9.0	375	2	Q98TX5_XENLA	Q98tx5	xenopus lae
27	100	9.0	574	2	Q8N4M9_HUMAN	Q8n4m9	homo sapien
28	98.5	8.8	375	2	Q5BL74_XENTR	Q5bl74	xenopus tro
29	97	8.7	379	1	CYR61_RAT	Q9es72	rattus norv
30	97	8.7	379	2	Q66HT5_RAT	Q66ht5	rattus norv
31	97	8.7	379	2	Q9WTM9_RAT	Q9wtm9	rattus norv
32	95.5	8.6	375	1	CEF10_CHICK	P19336	gallus gall
33	95	8.5	379	1	CYR61_MOUSE	P18406	mus musculu
34	95	8.5	379	2	Q3TX21_MOUSE	Q3tx21	m osteoclas
35	94.5	8.5	262	2	Q76C27_PAROL	Q76c27	paralichthy
36	93.5	8.4	988	2	Q97867_PIG	Q97867	sus scrofa
37	92.5	8.3	359	2	Q4SUZ2_TETNG	Q4suz2	tetraodon n
38	91.5	8.2	1410	2	Q20204_CAEEL	Q20204	caenorhabdi
39	91	8.2	381	2	Q53FA4_HUMAN	Q53fa4	homo sapien
40	91	8.2	498	2	Q5I092_XENTR	Q5i092	xenopus tro
41	90.5	8.1	299	2	Q7YRR8_PIG	Q7yrr8	sus scrofa
42	90	8.1	381	1	CYR61_HUMAN	O00622	homo sapien
43	90	8.1	381	2	Q6FI18_HUMAN	Q6fi18	homo sapien
44	89	8.0	984	2	Q67VW6_ORYSA	Q67vw6	oryza sativ
45	88.5	7.9	442	2	Q8CCV4_MOUSE	Q8ccv4	mus musculu

ALIGNMENTS

RESULT 1

SOSD1_HUMAN

ID SOSD1_HUMAN STANDARD; PRT; 206 AA.

AC Q6X4U4; Q96HJ7; Q9Y3U3;

DT 29-MAR-2005, integrated into UniProtKB/Swiss-Prot.

DT 29-MAR-2005, sequence version 2.

DT 07-MAR-2006, entry version 15.

DE Sclerostin domain-containing protein 1 precursor (Ectodermal BMP

DE inhibitor) (Ectodin) (Uterine sensitization-associated gene 1 protein)

DE (USAG-1).

GN Name=SOSTDC1; Synonyms=USAG1; ORFNames=CDA019;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=22984999; PubMed=14623234; DOI=10.1016/j.ydbio.2003.08.011;
RA Laurikkala J., Kassai Y., Pakkasjaervi L., Thesleff I., Itoh N.;
RT "Identification of a secreted BMP antagonist, ectodin, integrating
RT BMP, FGF, and SHH signals from the tooth enamel knot.";
RL Dev. Biol. 264:91-105 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Pheochromocytoma;
RA Liu F., Xu X.R., Qian B.Z., Xiao H., Chen Z., Han Z.;
RT "A novel gene expressed in human pheochromocytoma.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA O'Shaughnessy R.F.L., Yeo W., Gautier J., Jahoda C.A.B.,
RA Christiano A.M.;
RT "A novel secreted WNT agonist is a requirement for epithelial-
RT mesenchymal interactions.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,
RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,
RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
RA Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubbs K.,
RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA Waterston R.H., Wilson R.K.;
RT "The DNA sequence of human chromosome 7.";
RL Nature 424:157-164 (2003).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Bone marrow;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP PROTEIN SEQUENCE OF 24-38.
 RX PubMed=15340161; DOI=10.1110/ps.04682504;
 RA Zhang Z., Henzel W.J.;
 RT "Signal peptide prediction based on analysis of experimentally
 RT verified cleavage sites.";
 RL Protein Sci. 13:2819-2824(2004).
 RN [7]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 74-206.
 RC TISSUE=Brain;
 RG The German cDNA consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
 RN [8]
 RP FUNCTION, INTERACTIONS WITH BMP2; BMP4 AND BMP7, SUBCELLULAR LOCATION,
 RP AND TISSUE SPECIFICITY.
 RX PubMed=15020244; DOI=10.1016/j.bbrc.2004.02.075;
 RA Yanagita M., Oka M., Watabe T., Iguchi H., Niida A., Takahashi S.,
 RA Akiyama T., Miyazono K., Yanagisawa M., Sakurai T.;
 RT "USAG-1: a bone morphogenetic protein antagonist abundantly expressed
 RT in the kidney.";
 RL Biochem. Biophys. Res. Commun. 316:490-500(2004).
 CC -!- FUNCTION: May be involved in the onset of endometrial receptivity
 CC for implantation/sensitization for the decidual cell reaction
 CC Enhances Wnt signaling and inhibits TGF-beta signaling (By
 CC similarity). Directly antagonizes activity of BMP2, BMP4, BMP6 and
 CC BMP7 in a dose-dependent manner.
 CC -!- SUBUNIT: Interacts with BMP2, BMP4, BMP6 and BMP7 with high
 CC affinity.
 CC -!- SUBCELLULAR LOCATION: Secreted protein.
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and weakly in lung.
 CC -!- SIMILARITY: Belongs to the sclerostin family.
 CC -!- SIMILARITY: Contains 1 CTCK (C-terminal cystine knot-like) domain.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; AB059270; BAC20331.1; -; mRNA.
 DR EMBL; AF361494; AAL57219.1; -; mRNA.
 DR EMBL; AY255634; AAQ83296.1; -; mRNA.
 DR EMBL; AC079155; AAQ96855.1; -; Genomic_DNA.
 DR EMBL; BC008484; AAH08484.1; -; mRNA.
 DR EMBL; AL050024; CAB43243.2; -; mRNA.
 DR PIR; T08710; T08710.
 DR Ensembl; ENSG00000171243; Homo sapiens.
 DR HGNC; HGNC:21748; SOSTDC1.
 DR MIM; 609675; gene.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR008835; Sclerostin.
 DR PANTHER; PTHR14903; Sclerostin; 1.
 DR Pfam; PF05463; Sclerostin; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.

DR PROSITE; PS01225; CTCK_2; 1.
 KW Direct protein sequencing; Glycoprotein; Signal;
 KW Wnt signaling pathway.
 FT SIGNAL 1 23
 FT CHAIN 24 206 Sclerostin domain-containing protein 1.
 FT /FTId=PRO_0000033180.
 FT DOMAIN 75 170 CTCK.
 FT CARBOHYD 47 47 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 173 173 N-linked (GlcNAc . . .) (Potential).
 FT DISULFID 75 133 By similarity.
 FT DISULFID 89 147 By similarity.
 FT DISULFID 100 163 By similarity.
 FT DISULFID 104 165 By similarity.
 FT CONFLICT 8 14 FYLLPLA -> LSLIPLL (in Ref. 3).
 FT CONFLICT 20 20 S -> N (in Ref. 3).
 FT CONFLICT 41 41 P -> S (in Ref. 3).
 FT CONFLICT 62 62 N -> S (in Ref. 3).
 FT CONFLICT 69 69 T -> S (in Ref. 3).
 FT CONFLICT 106 106 P -> L (in Ref. 7).
 FT CONFLICT 127 127 S -> G (in Ref. 3).
 FT CONFLICT 182 182 M -> V (in Ref. 3).
 FT CONFLICT 188 188 V -> A (in Ref. 3).
 FT CONFLICT 205 205 M -> L (in Ref. 3).
 SQ SEQUENCE 206 AA; 23307 MW; 9FB3CC41E4B53834 CRC64;

Query Match 100.0%; Score 1115; DB 1; Length 206;
 Best Local Similarity 100.0%; Pred. No. 5.1e-100;
 Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPPAIHFYLLPLACILMKSCLA FKND EILYSHVVKPVP AHPSSNSTLNQARNGGRHF 60
 |||
 Db 1 MLPPAIHFYLLPLACILMKSCLA FKND EILYSHVVKPVP AHPSSNSTLNQARNGGRHF 60
 QY 61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
 |||
 Db 61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
 QY 121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180
 |||
 Db 121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180
 QY 181 SMSPAKPVQHHRRERKRASKSSKHSMS 206
 |||
 Db 181 SMSPAKPVQHHRRERKRASKSSKHSMS 206

RESULT 2

SOSD1_PONPY

ID SOSD1_PONPY STANDARD; PRT; 206 AA.
 AC Q5R5D2;
 DT 29-MAR-2005, integrated into UniProtKB/Swiss-Prot.
 DT 21-DEC-2004, sequence version 1.
 DT 07-FEB-2006, entry version 11.
 DE Sclerostin domain-containing protein 1 precursor.
 GN Name=SOSTDC1;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Pongo.
 OX NCBI_TaxID=9600;
 RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Kidney;
 RG The German cDNA consortium;
 RL Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Directly antagonizes activity of BMP2, BMP4, BMP6 and
 CC BMP7 in a dose-dependent manner. Enhances Wnt signaling and
 CC inhibits TGF-beta signaling. May be involved in the onset of
 CC endometrial receptivity for implantation/sensitization for the
 CC decidual cell reaction (By similarity).
 CC -!- SUBUNIT: Interacts with BMP2, BMP4, BMP6 and BMP7 with high
 CC affinity (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted protein (By similarity).
 CC -!- SIMILARITY: Belongs to the sclerostin family.
 CC -!- SIMILARITY: Contains 1 CTCK (C-terminal cystine knot-like) domain.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; CR860930; CAH93034.1; -; mRNA.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR008835; Sclerostin.
 DR PANTHER; PTHR14903; Sclerostin; 1.
 DR Pfam; PF05463; Sclerostin; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; 1.
 KW Glycoprotein; Signal; Wnt signaling pathway.
 FT SIGNAL 1 23 By similarity.
 FT CHAIN 24 206 Sclerostin domain-containing protein 1.
 FT /FTId=PRO_0000033182.
 FT DOMAIN 75 170 CTCK.
 FT CARBOHYD 47 47 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 173 173 N-linked (GlcNAc . . .) (Potential).
 FT DISULFID 75 133 By similarity.
 FT DISULFID 89 147 By similarity.
 FT DISULFID 100 163 By similarity.
 FT DISULFID 104 165 By similarity.
 SQ SEQUENCE 206 AA; 23297 MW; 0F8CCEBE364817E6 CRC64;

Query Match 99.3%; Score 1107; DB 1; Length 206;
 Best Local Similarity 99.5%; Pred. No. 3e-99;
 Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLPPAIHFYLLPLACILMKSCLA FKNDAT EILYSHVVKPVP AHPSSNSTLNQARNGGRHF 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MLPPAIHFYLLPLACILMKSCLA FKNDAT EILYSHVVKPVP AHPSSNSTLNQARNGGRHF 60
 Qy 61 SNTGLDRNTRVQVGCRELRSTKYI SDGQCTSI SPLKELVCAGECLPLVLPNWIGGGYGT 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 SNTGLDRNTRVQVGCRELRSTKYI SDGQCTSI SPLKELVCAGECLPLSVLPNWIGGGYGT 120
 Qy 121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180
 Qy 181 SMSPAKPVQHHRERKRASKSSKHSMS 206
 ||||||||||||||||||||||||||||
 Db 181 SMSPAKPVQHHRERKRASKSSKHSMS 206

RESULT 3
 SOSD1_MOUSE

ID SOSD1_MOUSE STANDARD; PRT; 206 AA.
 AC Q9CQN4; Q8CF09;
 DT 29-MAR-2005, integrated into UniProtKB/Swiss-Prot.
 DT 01-JUN-2001, sequence version 1.
 DT 07-MAR-2006, entry version 26.
 DE Sclerostin domain-containing protein 1 precursor (Ectodermal BMP
 DE inhibitor) (Ectodin) (Uterine sensitization-associated gene 1 protein)
 DE (USAG-1) (Sclerostin-like protein).
 GN Name=Sostdcl; Synonyms=Sostl, Usag1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], AND DEVELOPMENTAL STAGE.
 RC STRAIN=C57BL/6; TISSUE=Testis;
 RX MEDLINE=22505168; PubMed=12617826; DOI=10.1016/S1567-133X(02)00022-4;
 RA Menke D.B., Page D.C.;
 RT "Sexually dimorphic gene expression in the developing mouse gonad.";
 RL Gene Expr. Patterns 2:359-367(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, INTERACTIONS WITH BMP2; BMP4;
 RP BMP6 AND BMP7, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, DEVELOPMENTAL
 RP STAGE, AND INDUCTION.
 RX MEDLINE=22984999; PubMed=14623234; DOI=10.1016/j.ydbio.2003.08.011;
 RA Laurikkala J., Kassai Y., Pakkasjaervi L., Thesleff I., Itoh N.;
 RT "Identification of a secreted BMP antagonist, ectodin, integrating
 RT BMP, FGF, and SHH signals from the tooth enamel knot.";
 RL Dev. Biol. 264:91-105(2003).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RC STRAIN=C57BL/6;
 RA O'Shaughnessy R.F.L., Yeo W., Gautier J., Jahoda C.A.B.,
 RA Christiano A.M.;
 RT "A novel secreted WNT agonist is a requirement for epithelial-
 RT mesenchymal interactions.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6J; TISSUE=Kidney, and Pancreas;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Huminiecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,

RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavesi G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugiura K., Sultana R., Takenaka Y., Taki K.,
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RX PubMed=15020244; DOI=10.1016/j.bbrc.2004.02.075;
 RA Yanagita M., Oka M., Watabe T., Iguchi H., Niida A., Takahashi S.,
 RA Akiyama T., Miyazono K., Yanagisawa M., Sakurai T.;
 RT "USAG-1: a bone morphogenetic protein antagonist abundantly expressed
 RT in the kidney.";
 RL Biochem. Biophys. Res. Commun. 316:490-500(2004).
 CC -!- FUNCTION: May be involved in the onset of endometrial receptivity
 CC for implantation/sensitization for the decidual cell reaction.
 CC Enhances Wnt signaling and inhibits TGF-beta signaling (By
 CC similarity). Directly antagonizes activity of BMP2, BMP4, BMP6 and
 CC BMP7 in a dose-dependent manner.
 CC -!- SUBUNIT: Interacts with BMP2, BMP4, BMP6 and BMP7 with high
 CC affinity.
 CC -!- SUBCELLULAR LOCATION: Secreted protein.
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney at renal collecting

CC ducts level and weakly in brain.

CC -!- DEVELOPMENTAL STAGE: Expression was first detected at E11

CC throughout the surface of the embryo, and it was most intense in

CC the head region on the surfaces of the mandibular, maxillary, and

CC frontonasal processes. At E11.5 expression is detected in the

CC first and second branchial arches, pharynx and metanephros. At

CC E12-E14, expression was intense and strikingly confined to

CC developing ectodermal organs. The vibrissae, tylotrich hair

CC follicles, tongue papillae, and tooth germs as well as the ear

CC auricle. Also expressed intensely in kidney epithelium in the

CC stalk and tips of ureter as well as in the spermatid ducts in the

CC testis. At E17.5 strong expression was restricted to kidney

CC tubules and ameloblasts in teeth, and moderate expression was

CC observed in hair follicles, choroids plexus of the fourth cerebral

CC ventricle of the brain. First detected on E12.5 in interstitial

CC cell of the testis and increased towards E14.5. On 8 dpp (day post

CC partum) highly expression was detected in kidney and weakly in

CC skin.

CC -!- INDUCTION: Up-regulated by BMP2 and BMP7. Down-regulated by FGF4

CC and SHH.

CC -!- SIMILARITY: Belongs to the sclerostin family.

CC -!- SIMILARITY: Contains 1 CTCK (C-terminal cystine knot-like) domain.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; AY134666; AAN08617.1; -; mRNA.

DR EMBL; AB059271; BAC20332.1; -; mRNA.

DR EMBL; AY255635; AAQ83297.1; -; mRNA.

DR EMBL; AK002240; BAB21957.1; -; mRNA.

DR EMBL; AK002396; BAB22068.1; -; mRNA.

DR EMBL; AK007893; BAB25333.1; -; mRNA.

DR EMBL; AK007935; BAC25193.1; -; mRNA.

DR EMBL; AK007967; BAB25378.1; -; mRNA.

DR EMBL; BC021458; AAH21458.1; -; mRNA.

DR Ensembl; ENSMUSG00000036169; Mus musculus.

DR MGI; MGI:1913292; Sostdcl.

DR GO; GO:0005615; C:extracellular space; IDA.

DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0030514; P:negative regulation of BMP signaling pathway; IDA.

DR InterPro; IPR006207; Cys_knot_C.

DR InterPro; IPR008835; Sclerostin.

DR PANTHER; PTHR14903; Sclerostin; 1.

DR Pfam; PF05463; Sclerostin; 1.

DR PROSITE; PS01185; CTCK_1; FALSE_NEG.

DR PROSITE; PS01225; CTCK_2; 1.

KW Glycoprotein; Signal; Wnt signaling pathway.

FT	SIGNAL	1	23	By similarity.
FT	CHAIN	24	206	Sclerostin domain-containing protein 1.
FT				/FTId=PRO_0000033181.
FT	DOMAIN	75	170	CTCK.
FT	CARBOHYD	47	47	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	173	173	N-linked (GlcNAc. . .) (Potential).
FT	DISULFID	75	133	By similarity.
FT	DISULFID	89	147	By similarity.
FT	DISULFID	100	163	By similarity.
FT	DISULFID	104	165	By similarity.
FT	CONFLICT	69	69	S -> R (in Ref. 4; BAC25193).
FT	CONFLICT	80	80	S -> T (in Ref. 4; BAC25193).
SQ	SEQUENCE	206 AA;	23174 MW;	70D24819EB06CBC9 CRC64;

Query Match 95.7%; Score 1067; DB 1; Length 206;
 Best Local Similarity 94.7%; Pred. No. 2.4e-95;
 Matches 195; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```

QY      1 MLPPAIHFYLLPLACILMKSLAFKNDTEILYSHVVKPVAHPSSNSTLNQARNNGGRHF 60
      ||||| | : || ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MLPPAIHLSLIPLLCILMRNCLAFKNDTEILYSHVVKPVAHPSSNSTLNQARNNGGRHF 60

QY     61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVLPNWIGGGYGT 120
      ||||| : ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 SSTGLDRNSRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVLPNWIGGGYGT 120

QY    121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180

QY    181 SMSPAKPVQHHRERKRASKSSKHSMS 206
      | : ||||| ||||| ||||| ||||| : |
Db    181 SVSPAKPAQHHRERKRASKSSKHSLS 206

```

RESULT 4

SOSD1_RAT

ID SOSD1_RAT STANDARD; PRT; 206 AA.
 AC Q642G2; Q8CJA4;
 DT 29-MAR-2005, integrated into UniProtKB/Swiss-Prot.
 DT 25-OCT-2004, sequence version 1.
 DT 07-FEB-2006, entry version 13.
 DE Sclerostin domain-containing protein 1 precursor (Uterine
 DE sensitization-associated gene 1 protein) (USAG-1) (Wnt-signaling
 DE modulator).
 GN Name=Sostdcl; Synonyms=Usag1, Wise;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidea; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, TISSUE SPECIFICITY, AND
 RP INDUCTION.
 RC STRAIN=Sprague-Dawley; TISSUE=Uterus;
 RX MEDLINE=22277976; PubMed=12390898; DOI=10.1095/biolreprod.102.006858;
 RA Simmons D.G., Kennedy T.G.;
 RT "Uterine sensitization-associated gene-1: a novel gene induced within
 RT the rat endometrium at the time of uterine receptivity/sensitization
 RT for the decidual cell reaction."
 RL Biol. Reprod. 67:1638-1645(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Heart;
 RG NIH - Mammalian Gene Collection (MGC) project;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
 RP STAGE.
 RX PubMed=15373764; DOI=10.1111/j.0022-202X.2004.23410.x;
 RA O'Shaughnessy R.F.L., Yeo W., Gautier J., Jahoda C.A.B.,
 RA Christiano A.M.;
 RT "The WNT signalling modulator, Wise, is expressed in an interaction-
 RT dependent manner during hair-follicle cycling."
 RL J. Invest. Dermatol. 123:613-621(2004).

```

CC      -!- FUNCTION: Directly antagonizes activity of BMP2, BMP4, BMP6 and
CC      BMP7 in a dose-dependent manner (By similarity). May be involved
CC      in the onset of endometrial receptivity for
CC      implantation/sensitization for the decidual cell reaction.
CC      Enhances Wnt signaling and inhibits TGF-beta signaling.
CC      -!- SUBUNIT: Interacts with BMP2, BMP4, BMP6 and BMP7 with high
CC      affinity (By similarity).
CC      -!- SUBCELLULAR LOCATION: Secreted protein.
CC      -!- TISSUE SPECIFICITY: Highly expressed within the maximally
CC      sensitized/receptive endometrium. Weakly expressed in brain,
CC      kidney and the female reproductive tract. Expressed in the dermal
CC      papilla (DP) and at high level in the precortex of both anagen
CC      vibrissae and pelage follicles. Dynamic expression during the hair
CC      cycle.
CC      -!- DEVELOPMENTAL STAGE: Highly expressed in epidermis, dermis and the
CC      outermost periderm layer in the 17 day post-coitum (dpc).
CC      -!- INDUCTION: Up-regulated at day 5 pregnant or pseudopregnant of the
CC      uterine glandular epithelial cells, at time of maximal
CC      sensitization for the decidual cell reaction. Down-regulated at
CC      day 6 refractory uterus.
CC      -!- SIMILARITY: Belongs to the sclerostin family.
CC      -!- SIMILARITY: Contains 1 CTCK (C-terminal cystine knot-like) domain.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AF411056; AAN45848.1; -; mRNA.
DR      EMBL; BC081710; AAH81710.1; -; mRNA.
DR      Ensembl; ENSRNOG00000005770; Rattus norvegicus.
DR      RGD; 628877; Sostdcl.
DR      InterPro; IPR006207; Cys_knot_C.
DR      InterPro; IPR008835; Sclerostin.
DR      PANTHER; PTHR14903; Sclerostin; 1.
DR      Pfam; PF05463; Sclerostin; 1.
DR      PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR      PROSITE; PS01225; CTCK_2; 1.
KW      Glycoprotein; Signal; Wnt signaling pathway.
FT      SIGNAL          1      23      By similarity.
FT      CHAIN           24     206     Sclerostin domain-containing protein 1.
FT                                     /FTId=PRO_0000033183.
FT      DOMAIN          75     170     CTCK.
FT      CARBOHYD         47      47     N-linked (GlcNAc. . .) (Potential).
FT      CARBOHYD        173     173     N-linked (GlcNAc. . .) (Potential).
FT      DISULFID         75     133     By similarity.
FT      DISULFID         89     147     By similarity.
FT      DISULFID        100     163     By similarity.
FT      DISULFID        104     165     By similarity.
FT      CONFLICT        187     187     P -> L (in Ref. 1).
SQ      SEQUENCE        206 AA;  23136 MW;  52BD602431C6DD15 CRC64;

```

Query Match 95.2%; Score 1062; DB 1; Length 206;
Best Local Similarity 94.7%; Pred. No. 7.2e-95;
Matches 195; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MLPPAIHFYLLPLACILMKSCIAFKNDATEILYSHVVKPVPAPHPSSNSTLNQARNGGRHF 60
| | | | | : | | | | : | | | | | | | | | | | | | | | | | | |
Db 1 MLPPAIHLSLIPLLCILMKNCLAFKNDATEILYSHVVKPVSAHPSSNSTLNQARNGGRHF 60

Qy 61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVLPNWIGGGYGT 120
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 SSTGLDRNSRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVLPNWIGGGYGT 120

```

Qy      121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180
        |||
Db      121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180

Qy      181 SMSPAKPVQHHRRERKRASKSSKHSMS 206
        |:|||||
Db      181 SVSPAQPAQHHRRERKRASKSSKHSLS 206

```

RESULT 5

SOSD1_CHICK

```

ID  SOSD1_CHICK      STANDARD;          PRT;    206 AA.
AC  Q6VYA3;
DT  29-MAR-2005, integrated into UniProtKB/Swiss-Prot.
DT  05-JUL-2004, sequence version 1.
DT  07-FEB-2006, entry version 10.
DE  Sclerostin domain-containing protein 1 precursor (Wnt-signaling
DE  modulator).
GN  Name=SOSTDC1; Synonyms=WISE;
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, INTERACTION WITH LRP6, AND
RP  DEVELOPMENTAL STAGE.
RX  MEDLINE=22781984; PubMed=12900447; DOI=10.1242/dev.00674;
RA  Itasaki N., Jones C.M., Mercurio S., Rowe A., Domingos P.M.,
RA  Smith J.C., Krumlauf R.;
RT  "Wise, a context-dependent activator and inhibitor of Wnt
RT  signalling.";
RL  Development 130:4295-4305(2003).
CC  -!- FUNCTION: Can activate or inhibit Wnt signaling in a context-
CC  dependent manner. Activates the canonical Wnt pathway whereby acts
CC  through Disheveled proteins and beta-catenin. Antagonises Wnt
CC  signaling through the canonical pathways presumably by blocking
CC  accessibility of certain WNTs to their receptors. Induces posterior
CC  neural markers via components of the canonical Wnt pathway.
CC  -!- SUBUNIT: Interacts with LRP6.
CC  -!- SUBCELLULAR LOCATION: Secreted protein.
CC  -!- DEVELOPMENTAL STAGE: Expression was first detected broadly at
CC  stage 9, and then localised in the posterior surface ectoderm
CC  overlying the presomitic mesoderm by stage 10-11 embryo.
CC  -!- SIMILARITY: Belongs to the sclerostin family.
CC  -!- SIMILARITY: Contains 1 CTCK (C-terminal cystine knot-like) domain.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
DR  EMBL; AY319926; AAQ57626.1; -; mRNA.
DR  InterPro; IPR006207; Cys_knot_C.
DR  InterPro; IPR008835; Sclerostin.
DR  PANTHER; PTHR14903; Sclerostin; 1.
DR  Pfam; PF05463; Sclerostin; 1.
DR  PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR  PROSITE; PS01225; CTCK_2; 1.
KW  Glycoprotein; Signal; Wnt signaling pathway.
FT  SIGNAL      1      22      Potential.
FT  CHAIN       23     206      Sclerostin domain-containing protein 1.

```

FT				/FTId=PRO_0000033184.
FT	DOMAIN	75	170	CTCK.
FT	CARBOHYD	47	47	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	173	173	N-linked (GlcNAc. . .) (Potential).
FT	DISULFID	75	133	By similarity.
FT	DISULFID	89	147	By similarity.
FT	DISULFID	100	163	By similarity.
FT	DISULFID	104	165	By similarity.
SQ	SEQUENCE	206 AA;	23144 MW;	75F438854599CEAD CRC64;

Query Match 86.9%; Score 969; DB 1; Length 206;
Best Local Similarity 87.9%; Pred. No. 8e-86;
Matches 181; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

Qy	1	MLPPAIHFYLLPLACILMKSCIAFKNDATEILYSHVVKPVAHPSSNSTLNQARNGGRHF	60
		: :	
Db	1	MLLSAIHFYGLLLACTFTRSYSAFKNDATEILYSHVVKPAPASPSSNSTLNQARNGGRHY	60
Qy	61	SNTGLDRNTRVQVGCRELRSTKYISDGQCTSISPLKELVCAGECLPLPVLPNWIGGGYGT	120
		: :	
Db	61	AGTGSDRNNRVQVGCRELRSTKYISDGQCTSINPLKELVCAGECLPLPLLPNWIGGGYGT	120
Qy	121	KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRYKITVVTACKCKRYTRQHNESSHNFE	180
Db	121	KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSIRTYKITVVTACKCKRYTRQHNESSHNFE	180
Qy	181	SMSPAKPVQHHRERKRASKSSKHSMS	206
		:	
Db	181	GTSQAKPVQHHKERKRASKSSKHSTS	206

RESULT 6

```

Q6VYA2_XENLA
ID   Q6VYA2_XENLA    PRELIMINARY;   PRT;   213 AA.
AC   Q6VYA2;
DT   05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT   05-JUL-2004, sequence version 1.
DT   07-FEB-2006, entry version 9.
DE   Wise-A.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipodea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22781984; PubMed=12900447; DOI=10.1242/dev.00674;
RA   Itasaki N., Jones C.M., Mercurio S., Rowe A., Domingos P.M.,
RA   Smith J.C., Krumlauf R.;
RT   "Wise, a context-dependent activator and inhibitor of Wnt
RT   signalling.";
RL   Development 130:4295-4305(2003).
CC   -----
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; AY319927; AAQ57627.1; -; mRNA.
DR   InterPro; IPR006207; Cys_knot_C.
DR   InterPro; IPR008835; Sclerostin.
DR   PANTHER; PTHR14903; Sclerostin; 1.
DR   Pfam; PF05463; Sclerostin; 1.

```

DR PROSITE; PS01225; CTCK_2; 1.
SQ SEQUENCE 213 AA; 24359 MW; 3272FC5357A4701D CRC64;

Query Match 73.3%; Score 817.5; DB 2; Length 213;
Best Local Similarity 75.2%; Pred. No. 4.6e-71;
Matches 152; Conservative 21; Mismatches 20; Indels 9; Gaps 3;

Qy 10 LLPLACILMKSCLA FKNDATEILYSHVVKPVPAPHPSSNSTLNQARNGGRHFSNTGLDRNT 69
:| |||||::||::||| ||||| ||| : :::||| ||||| ||| :|: |||
Db 10 MLYLACILIESCVSFKNDATEILYSHVDKHI-QDSANSSTLNQARNGGRNAANSALDRTN 68

Qy 70 RVQVGCRELRSTKYISDGQCTSISPLKELVCAGECLPLPVLPNWIGGGYGTKYWSRRSSQ 129
||| ||||| ||||| ||||| ||||| ||||| :||| ||||| ||||| ||||| |||||
Db 69 HHQVGCRELRSTKYISDGQCTSIQPLKELVCAGECLPLPILPNWIGGGYGLKYWSRRSSQ 128

Qy 130 EWRCVNDKTRTQRIQLQCQDGSTRYKITVVTACKCKRYTRQHNESSHNFSMSPAKPV- 188
||| ||||| ||||| ||||| :||: ||||: ||||: ||||| ||||| ||||| :||| |||
Db 129 EWRCVNDKTRTQRIQLQCEDGTTTRYKVTVVTSCCKRYTRQHNESSHNYQGASPIKPVH 188

Qy 189 --QHH-----RERKRASKSSKH 203
||| ||: || |||
Db 189 SHQHHPHHNRDKKRLIKMSKH 210

RESULT 7

Q688C5_XENTR

ID Q688C5_XENTR PRELIMINARY; PRT; 213 AA.

AC Q688C5;

DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.

DT 11-OCT-2004, sequence version 1.

DT 07-FEB-2006, entry version 10.

DE Wise-A.

OS *Xenopus tropicalis* (Western clawed frog) (*Silurana tropicalis*).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

OC Xenopodinae; *Xenopus*; *Silurana*.

OX NCBI_TaxID=8364;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Qin S., Dors M., Johnson E., Bloom S., Hood L., Rowen L.;

RT "Sequence of *xenopus tropicalis* development genes.";

RL Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC

DR EMBL; AC151465; AAU12851.1; -; Genomic_DNA.

DR Ensembl; ENSXETG00000022798; *Xenopus tropicalis*.

DR InterPro; IPR006207; Cys_knot_C.

DR InterPro; IPR008835; Sclerostin.

DR PANTHER; PTHR14903; Sclerostin; 1.

DR Pfam; PF05463; Sclerostin; 1.

DR PROSITE; PS01225; CTCK_2; 1.

SQ SEQUENCE 213 AA; 24400 MW; 801813CF69A10AC2 CRC64;

Query Match 72.8%; Score 811.5; DB 2; Length 213;
Best Local Similarity 73.8%; Pred. No. 1.7e-70;
Matches 149; Conservative 21; Mismatches 23; Indels 9; Gaps 3;

Qy 10 LLPLACILMKSCLA FKNDATEILYSHVVKPVPAPHPSSNSTLNQARNGGRHFSNTGLDRNT 69
:| ||| :||: ||||| ||||| ||| : :::| ||||| ||||| :|: :||

```

Db          10 MLYFACIFIESCMSFKNDATILYSHVDKNI-QESANSSALNQARNGGRHTANSAMDRTN   68
Qy          70 RVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVLPNWIGGGYGTKYWSRRSSQ  129
           |||||
Db          69 PHQVGCRELRSTKYISDGQCTSIQLKELVCAGECLPLPILPNWIGGGYGLKYWSRRSSQ  128
Qy          130 EWRCVNDKTRTQRIQLQCQDGSTRTYKITVVVTACKCKRYTRQHNESSHNFESMSPAKPV-  188
           |||||
Db          129 EWRCVNDKTRTQRIQLQCEDGTTRTYKVTVVTSCKCKRYTRQHNESSHNYEGASPMKPIH  188
Qy          189 --QHH-----RERKRASKSSKH  203
           |||      |:| | |||
Db          189 SLOHHHSHHNDRDKRLIKMSKH  210

```

RESULT 8

```
Q6X4U2_XENLA
ID   Q6X4U2_XENLA    PRELIMINARY;   PRT;   213 AA.
AC   Q6X4U2;
DT   05-JUL-2004, integrated into UniProtKB/TREMBL.
DT   05-JUL-2004, sequence version 1.
DT   07-FEB-2006, entry version 11.
DE   Uterine sensitization-associated protein-1.
GN   Name=Usag1;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   O'Shaughnessy R.F.L., Yeo W., Gautier J., Jahoda C.A.B.,
RA   Christiano A.M.;
RT   "A novel secreted WNT agonist is a requirement for epithelial-
RT   mesenchymal interactions.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -----
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; AY255636; AAQ83298.1; -; mRNA.
DR   InterPro; IPR006207; Cys_knot_C.
DR   InterPro; IPR008835; Sclerostin.
DR   PANTHER; PTHR14903; Sclerostin; 1.
DR   Pfam; PF05463; Sclerostin; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
SQ   SEQUENCE    213 AA;  24286 MW;  43E0AFCF6A78FFC5 CRC64;
```

Query Match 71.4%; Score 796.5; DB 2; Length 213;
Best Local Similarity 73.8%; Pred. No. 5e-69;
Matches 149; Conservative 22; Mismatches 22; Indels 9; Gaps 3;

Qy	10	LLPLACILMKSCLA FKNDAT EILY SHVVKPVPAPHPSSNSTLNQARNGGRHFSNTGLDRNT	69
		: :: : : :: : :	
Db	10	MLYLACILIESCVSFKNDAT EILY SHVDKHI-QDSANSSTLNQARNGGRNAANSALDRNT	68
Qy	70	RVQVGCRELRSTKYISDGQCT SISPLKELVCAGECLPLPVLPNWIGGGYGTKYWSRRSSQ	129
		: :	
Db	69	HHQVGCRELRSTKYISDGQCT SIQPLKELVCAGECLPLSILAHWIGGGYGLKYWSRRSSQ	128
Qy	130	EWRCVNDKTRTORIQLQCQDGSTR TYKITVVTACKCKRYTROHNES SHNFESMSPAKPV-	188


```

CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; BC078598; AAH78598.1; -; mRNA.
DR      InterPro; IPR006207; Cys_knot_C.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR008835; Sclerostin.
DR      PANTHER; PTHR14903; Sclerostin; 1.
DR      Pfam; PF05463; Sclerostin; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PROSITE; PS01225; CTCK_2; 1.
KW      Hypothetical protein.
FT      NON_TER      1      1
SQ      SEQUENCE      187 AA; 21320 MW; 4670BF6C3F8D741A CRC64;

Query Match      66.5%; Score 741.5; DB 2; Length 187;
Best Local Similarity 75.5%; Pred. No. 9.6e-64;
Matches 139; Conservative 15; Mismatches 21; Indels 9; Gaps 3;

Qy      28 ATEILYSHVVKPVPAPHPSSNSTLNQARNGGRHFSNTGLDRNTRVQVGCRELRSTKYISDG 87
      ||||| | | ::||| ||||| : |: ||| ||||| |||||
Db      2 ATEILYSQVDKHT-QESANSSTLNQARNGGRNTANSALDRTNHHQVGCRELRSTKYISDG 60

Qy      88 QCTSISPLKELVCAGECLPLPVLPNWIGGGYGTKYWSRRSSQEWRCVNDKTRTQRIQLQC 147
      |||| | ||||| ||||| : ||||| ||||| ||||| ||||| ||||| |||||
Db      61 QCTSIQPLKELVCAGECLPLPILPNWIGGGYGLKYWSRRSSQEWRCVNDKTRTQRIQLQC 120

Qy      148 QDGSTRTYKITVVTACKCKRYTRQHNESSHNFE SMSPAKPV---QHH-----RERKRASK 199
      :||: ||||: |||: ||||| ||||| ||||| : || |||: ||| :||| |
Db      121 EDGTTTRYKVTVVTSCKCKRYTRQHNESSHNYEGASPIKPIHSQQHHHSHHNDRDKKRLIK 180

Qy      200 SSKH 203
      |||
Db      181 MSKH 184

```

RESULT 10

```

Q568H6_BRARE
ID   Q568H6_BRARE    PRELIMINARY;   PRT;   207 AA.
AC   Q568H6;
DT   10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT   10-MAY-2005, sequence version 1.
DT   07-FEB-2006, entry version 4.
DE   Hypothetical protein zgc:110293.
GN   ORFNames=zgc:110293;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Olfactory epithelium;
 RG NIH MGC Project;
 RL Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC092855; AAH92855.1; -; mRNA.
 DR Ensembl; ENSDARG00000026227; Danio rerio.
 DR ZFIN; ZDB-GENE-050417-64; zgc:110293.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR008835; Sclerostin.
 DR PANTHER; PTHR14903; Sclerostin; 1.
 DR Pfam; PF05463; Sclerostin; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 207 AA; 23503 MW; 071B9F3C99A8D8EB CRC64;

Query Match 57.8%; Score 645; DB 2; Length 207;
 Best Local Similarity 64.7%; Pred. No. 2.6e-54;
 Matches 130; Conservative 20; Mismatches 37; Indels 14; Gaps 6;

QY 3 PPAIHFYLLPLACILMKSLAFKNDAT EILYSHVVKPVAHPSSNSTLNQARNGGRHFSN 62
 | : : | : | | : : | | | | | | | | | | | : : | : | : | | | |
 Db 6 PESCNF--MVLFCFLIRSGTLTKNDAT EIFYSHVVSPPV-QDAQSNASLNRRASGGRGFST 62

QY 63 TGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVLPNWIGGGYGTKY 122
 | | | : | | | | | | | | | | | | | : : | | | : | | | | | | |
 Db 63 H--DRE-RIPVGCRELRSTKYISDGQCTSI NPKGLVCTGQCLPAQMLPNWI-GGYGKKS 118

QY 123 WSRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESHNFSM 182
 | : | : | | | | | | | | | | | | | | | | | : | | | | : | | | | |
 Db 119 WNRNSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTSCCKRYSRQHNESGVKSEGY 178

QY 183 SPAKPVQHHRRERKRASKSSKH 203
 | : : | | |
 Db 179 S-----HSQIKKTEKQSGH 192

RESULT 11

Q6VYA1_XENLA

ID Q6VYA1_XENLA PRELIMINARY; PRT; 153 AA.
 AC Q6VYA1;
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 07-FEB-2006, entry version 7.
 DE Wise-B (Fragment).

Query Match 55.2%; Score 615.5; DB 2; Length 153;
Best Local Similarity 74.7%; Pred. No. 1.4e-51;
Matches 115; Conservative 15; Mismatches 23; Indels 1; Gaps 1;

RESULT 12

```
ID      Q4RYK3_TETNG    PRELIMINARY;    PRT;    179 AA.
AC      Q4RYK3;
DT      19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT      19-JUL-2005, sequence version 1.
DT      07-FEB-2006, entry version 4.
DE      Chromosome 2 SCAF14976, whole genome shotgun sequence. (Fragment).
GN      ORFNames=GSTENG00026927001;
OS      Tetraodon nigroviridis (Green puffer).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC      Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC      Tetradontoidea; Tetraodontidae; Tetraodon.
OX      NCBI_TaxID=99883;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      PubMed=15496914; DOI=10.1038/nature03025;
RA      Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA      Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
```

RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; CAAE01014976; CAG06529.1; -; Genomic_DNA.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR008835; Sclerostin.
 DR PANTHER; PTHR14903; Sclerostin; 1.
 DR Pfam; PF05463; Sclerostin; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 FT NON_TER 179 179
 SQ SEQUENCE 179 AA; 19644 MW; C569E31E642A2D25 CRC64;

Query Match 33.4%; Score 372.5; DB 2; Length 179;
 Best Local Similarity 44.6%; Pred. No. 7.1e-28;
 Matches 78; Conservative 31; Mismatches 45; Indels 21; Gaps 6;

QY 14 ACILMKSCS-----AFKNDATEILYS-----HVVKPVPAHPSSNSTLNQARNGGRHF 60
 | :|::| | | | | | : | | | : | :|::| |
 Db 13 ALVLLQGCCAAARGWKALKNDATVLAADDRERSPEPAHAANASSNAG-NRAKSGARST 71
 QY 61 SNTGLDRNTRVQVGCRELSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
 | : : : | | | | | : : : | | | : | | | : : | | |
 Db 72 STVSYSAS---ELSCRELSTRYVTDGSCRSAPKIKELVCSGQCLPAHLMPTIG---RA 125
 QY 121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRYKITVVTACKCKRYTRQHNS 175
 | : | | : : | | : : | | | : : | | | | : : | | : : | |
 Db 126 KWW-RSSTSEYRCVPAHSRPRRIQLRCPNGNTRTYKIRTVTSCKCKRFRAHNNQS 179

RESULT 13

SOST_HUMAN

ID SOST_HUMAN STANDARD; PRT; 213 AA.
 AC Q9BQB4;
 DT 05-MAR-2002, integrated into UniProtKB/Swiss-Prot.
 DT 01-JUN-2001, sequence version 1.
 DT 07-MAR-2006, entry version 38.
 DE Sclerostin precursor.

GN Name=SOST; ORFNames=UNQ2976/PRO7455/PRO7476;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RX PubMed=11181578; DOI=10.1093/hmg/10.5.537;
 RA Balemans W., Ebeling M., Patel N., van Hul E., Olson P., Dioszegi M.,
 RA Lacza C., Wuyts W., van den Ende J., Willems P., Paes-Alves A.F.,
 RA Hill S., Bueno M., Ramos F.J., Tacconi P., Dikkers F.G., Stratakis C.,
 RA Lindpaintner K., Vickery B., Foernzler D., Van Hul W.;
 RT "Increased bone density in sclerosteosis is due to the deficiency of a
 RT novel secreted protein (SOST).";
 RL Hum. Mol. Genet. 10:537-543 (2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alish R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cystine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589 (2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
 RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
 RA Wood W.I., Godowski P.J., Gray A.M.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270 (2003).
 RN [4]
 RP PROTEIN SEQUENCE OF 24-38.
 RX PubMed=15340161; DOI=10.1110/ps.04682504;
 RA Zhang Z., Henzel W.J.;
 RT "Signal peptide prediction based on analysis of experimentally
 RT verified cleavage sites.";
 RL Protein Sci. 13:2819-2824 (2004).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis.
 CC -!- SUBCELLULAR LOCATION: Secreted protein (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9BQB4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9BQB4-2; Sequence=VSP_010189;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Widely expressed at low levels with highest
 CC levels in bone, cartilage, kidney, liver, bone marrow and primary
 CC osteoblasts differentiated for 21 days.

```

CC      -!- DISEASE: Defects in SOST are the cause of sclerosteosis
CC          [MIM:269500]. Sclerosteosis is an autosomal recessive sclerosing
CC          bone dysplasia characterized by progressive skeletal overgrowth.
CC          The disorder is similar to van Buchem hyperostosis corticalis
CC          generalisata [MIM:239100], but differing in radiologic appearance
CC          of the bone changes and in the presence of asymmetric cutaneous
CC          syndactyly of the index and middle fingers in many cases. The
CC          majority of affected individuals have been reported in the
CC          Afrikaner population of South Africa, where a high incidence of
CC          the disorder occurs as a result of a founder effect.
CC      -!- SIMILARITY: Belongs to the sclerostin family.
CC      -!- SIMILARITY: Contains 1 CTCK (C-terminal cystine knot-like) domain.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AF331844; AAK16158.1; -; mRNA.
DR      EMBL; AF326736; AAK13451.1; -; Genomic_DNA.
DR      EMBL; AF326739; AAK13454.1; -; mRNA.
DR      EMBL; AY358203; AAQ88570.1; -; mRNA.
DR      EMBL; AY358627; AAQ88990.1; -; mRNA.
DR      Ensembl; ENSG00000167941; Homo sapiens.
DR      HGNC; HGNC:13771; SOST.
DR      MIM; 269500; phenotype.
DR      MIM; 605740; gene.
DR      GO; GO:0030279; P:negative regulation of ossification; NAS.
DR      InterPro; IPR006207; Cys_knot_C.
DR      InterPro; IPR008835; Sclerostin.
DR      PANTHER; PTHR14903; Sclerostin; 1.
DR      Pfam; PF05463; Sclerostin; 1.
DR      PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR      PROSITE; PS01225; CTCK_2; FALSE_NEG.
KW      Alternative splicing; Direct protein sequencing; Glycoprotein; Signal.
FT      SIGNAL          1          23
FT      CHAIN           24         213          Sclerostin.
FT                                     /FTid=PRO_0000033177.
FT      DOMAIN          82         172          CTCK.
FT      CARBOHYD         53         53          N-linked (GlcNAc. . .) (Potential).
FT      CARBOHYD        175        175          N-linked (GlcNAc. . .) (Potential).
FT      VARSPLIC         64         73          RPPHHPFETK -> WPGGRPPSRAPLST (in isoform
FT                                     2).
FT                                     /FTid=VSP_010189.
SQ      SEQUENCE      213 AA;  24031 MW;  30DBD55CE73D5BB2 CRC64;

```

Query Match 31.3%; Score 349.5; DB 1; Length 213;
Best Local Similarity 40.5%; Pred. No. 1.5e-25;
Matches 85; Conservative 33; Mismatches 65; Indels 27; Gaps 10;

Qy	11	LPLA---CILMKSCS-----AFKNDATEIL--YSHVVKPVAHPSSNSTLNQARN	55
		: : : : : : :	
Db	3	LPLALCLVCLLVHTAFRVVEGQGWQAFKNDATEIIPELGEYPEP-PPELENNKTMNRAEN	61
Qy	56	GGR--HFSNTGLDRNTRVQVGCRELIRSTKYISDGQCTSSISPLKELVCAGECLPLPVLPNW	113
		: : : : : : :	
Db	62	GGRPPHHPFETKDVS---EYSCRELHFTRYVTDGPCRSAPVTELVCSGQCGPARLLPNA	118
Qy	114	IGGGYGTKYWSRRSSQEWRCVNDKTRTORIQLQCQDG-STRTYKITVVTACKCKRYTRQH	172
		: : : : : : : : :	
Db	119	IGRG---KWW-RPSGPDFRCIPDRYRAQRVQLLCPGGEAPRARKVRLVASCKCKRLTRFH	174
Qy	173	NESSHNFEISMSPAKPVOHHRERKRASKSSK	202

Db 175 NQSELKDFGTEAARPQKGRKPRPRA-RSAK 203

RESULT 14

Q495N9_HUMAN

ID Q495N9_HUMAN PRELIMINARY; PRT; 213 AA.
AC Q495N9;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Sclerostin,.
GN Name=SOST;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (AUG-2005) to the EMBL/GenBank/DDBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC101087; AAI01088.1; -; mRNA.
DR EMBL; BC101088; AAI01089.1; -; mRNA.
DR EMBL; BC101086; AAI01087.1; -; mRNA.
DR EMBL; BC101089; AAI01090.1; -; mRNA.
DR InterPro; IPR006207; Cys_knot_C.
DR Pfam; PF05463; Sclerostin; 1.
SQ SEQUENCE 213 AA; 24031 MW; 30DBD55CE73D5BB2 CRC64;

Query Match 31.3%; Score 349.5; DB 2; Length 213;
Best Local Similarity 40.5%; Pred. No. 1.5e-25;

Matches 85; Conservative 33; Mismatches 65; Indels 27; Gaps 10;

```

Qy      11 LPLA---CILMKSC-----AFKNDATEIL--YSHVVKPVPAPHPSSNSTLNQARN 55
      ||| |   |:| :   :   ||||| | | | :   :| |   :| |:| :| |
Db      3 LPLALCLVCLLVHTAFRVVEGQGWQAFKNDATEIIPELGEYPEP-PPELENNKTMNRAEN 61

Qy     56 GGR--HFSNTGLDRNTRVQVGCRELIRSTKYISDGQCTSI SPLKELVCAGECLPLPVLPNW 113
      ||| |   | :   : |||| | |:| :| | | | : |||| |:| | | : |||
Db     62 GGRPPHHPFETKDVS---EYSCRELHFTRYVTDGPCRSAKPVTELVCSGQC GPARLLPNA 118

Qy    114 IGGYGTYKYWSRRSSQEWRVCVNDKTRTQRIQLQCQDG-STRTYKITVVTACKCKRYTRQH 172
      || |   |:| | | :| |:| :| | | | | : | | : :| :||| | | |
Db    119 IGRG---KWW-RPSGPDFRCIPDRYRAQRVQLLCPGGEAPRARKVRLVASCKCKRLTRFH 174

Qy    173 NESSHNFESMSPAKPVQHHRRERKRASKSSK 202
      |:|   |:| :   : | | | |:| |
Db    175 NQSELKDFGTEAARPQKGRKPRPRA-RSAK 203

```

RESULT 15

SOST_CERAE

```

ID  SOST_CERAE      STANDARD;      PRT;      213 AA.
AC  Q9BG78;
DT  05-MAR-2002, integrated into UniProtKB/Swiss-Prot.
DT  01-JUN-2001, sequence version 1.
DT  07-FEB-2006, entry version 26.
DE  Sclerostin precursor.
GN  Name=SOST;
OS  Cercopithecus aethiops (Green monkey) (Grivet).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC  Cercopithecidae; Cercopithecinae; Cercopithecus.
OX  NCBI_TaxID=9534;
RN  [1]
RP  NUCLEOTIDE SEQUENCE [MRNA].
RX  MEDLINE=21090529; PubMed=11179006;
RA  Brunkow M.E., Gardner J.C., Van Ness J., Paeper B.W., Kovacevich B.R.,
RA  Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA  Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA  Beighton P., Mulligan J.T.;
RT  "Bone dysplasia sclerosteosis results from loss of the SOST gene
RT  product, a novel cystine knot-containing protein.";
RL  Am. J. Hum. Genet. 68:577-589(2001).
CC  -!- FUNCTION: Seems to play a role in bone homeostasis (By
CC  similarity).
CC  -!- SUBCELLULAR LOCATION: Secreted protein (Potential).
CC  -!- SIMILARITY: Belongs to the sclerostin family.
CC  -!- SIMILARITY: Contains 1 CTCK (C-terminal cystine knot-like) domain.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
DR  EMBL; AF326742; AAK13457.1; -; mRNA.
DR  InterPro; IPR006207; Cys_knot_C.
DR  InterPro; IPR008835; Sclerostin.
DR  PANTHER; PTHR14903; Sclerostin; 1.
DR  Pfam; PF05463; Sclerostin; 1.
DR  PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR  PROSITE; PS01225; CTCK_2; FALSE_NEG.
KW  Glycoprotein; Signal.
FT  SIGNAL      1      23      Potential.

```


FT CHAIN 24 213 Sclerostin.
 FT /FTId=PRO_0000033176.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 175 175 N-linked (GlcNAc. . .) (Potential).
 SQ SEQUENCE 213 AA; 23908 MW; 6DA7B5EDE674728A CRC64;

Query Match 31.2%; Score 348; DB 1; Length 213;
 Best Local Similarity 40.5%; Pred. No. 2.1e-25;
 Matches 83; Conservative 31; Mismatches 65; Indels 26; Gaps 9;

Qy 11 LPLA---CILMKSCS-----AFKNDATEIL--YSHVVKPVPAPHPSSNSTLNQARN 55
 |||| |::: | |||||::: :| | :| |::: |
 Db 3 LPLALCLVCLLVHAAFRVVEGQGWQAFKNDATEIIPELGEYPEP-PPELENNKTMNRAEN 61

 Qy 56 GGR--HFSNTGLDRNTRVQVGCRELRSTKYISDGQCTSISPLKELVCAGECLPLPVLPNW 113
 ||| | : : |||| |:::| | | : ||||:|:| | :|||
 Db 62 GGRPPHHPFETKDVS---EYSCRELHFTRYVTDGPCRSAKPVTELVCSGQCGPARLLPNA 118

 Qy 114 IGGGYGTKYWSRRSSQEWRCVNDKTRTQRIQLQCQDGST-RTYKITVVTACKCKRYTRQH 172
 || | :| | | ::||: | : | ||:| | | : | | : :| :||| | | |
 Db 119 IGRG---KWW-RPSGPDFRCIPDRYRAQRVQLLCPGGAAPRARKVRLVASCKCKRLTRFH 174

 Qy 173 NESSHNFESMSPAKPVQHHREKRKA 197
 |:| |:| : : |||
 Db 175 NQSELKDFGPEAAPQKGRKPRPRA 199

Search completed: August 12, 2006, 07:00:54

Job time : 303 secs

SCORE 1.3 BuildDate: 12/06/2005

SCORE Search Results Details for Application 10627604 and Search Result us-10-627-604-2.rag.

[Score Home](#)
[Retrieve Application](#)
[SCORE System](#)
[SCORE](#)
[Comments /](#)
[Page](#)
[List](#)
[Overview](#)
[FAQ](#)
[Suggestions](#)

This page gives you Search Results detail for the Application 10627604 and Search Result us-10-627-604-2.rag.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2006, 06:52:05 ; Search time 196 Seconds
(without alignments)
480.544 Million cell updates/sec

Title: US-10-627-604-2
Perfect score: 1115
Sequence: 1 MLPPAIHFYLLPLACILMKS.....PVQHHRRERKRASKSSKHSMS 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_8:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1115	100.0	206	2	AAW09408	Aaw09408 Human sma
2	1115	100.0	206	2	AAW95711	Aaw95711 Homo sapi
3	1115	100.0	206	3	AAB10233	Aab10233 Human adu
4	1115	100.0	206	5	ABB84851	Abb84851 Human PRO
5	1115	100.0	206	5	ABB09210	Abb09210 SCGF CNN
6	1115	100.0	206	5	ABB95457	Abb95457 Human ang
7	1115	100.0	206	7	ADD10359	Add10359 Human sec
8	1115	100.0	206	7	ADD11319	Add11319 Human sec
9	1115	100.0	206	7	ADD37112	Add37112 Human sec
10	1115	100.0	206	8	ADE41320	Ade41320 Human sec
11	1115	100.0	206	8	ADH43503	Adh43503 Human PRO
12	1115	100.0	206	8	ADI27203	Adi27203 Human LRP
13	1115	100.0	206	8	ADK70477	Adk70477 Respirato
14	1115	100.0	206	8	ADK82848	Adk82848 Human PRO
15	1115	100.0	206	8	ADQ26750	Adq26750 Form proc
16	1115	100.0	206	9	ADX07493	Adx07493 Cyclin-de
17	1115	100.0	206	9	ADY74839	Ady74839 Human par
18	1109	99.5	206	2	AAW58704	Aaw58704 Human sma
19	1109	99.5	206	5	AAU82653	Aau82653 Human bre
20	1109	99.5	206	5	ABB09200	Abb09200 Human sma
21	1088	97.6	230	7	ADB65448	Adb65448 Human pro
22	1067	95.7	206	3	AAY76031	Aay76031 Murine sk
23	1067	95.7	206	3	AAY75981	Aay75981 Murine sk
24	1067	95.7	206	4	AAB55920	Aab55920 Skin cell
25	1067	95.7	206	4	AAB55970	Aab55970 Skin cell
26	1067	95.7	206	5	ABB72170	Abb72170 Murine pr
27	1067	95.7	206	5	ABB72120	Abb72120 Murine pr
28	1067	95.7	206	8	ADS85127	Ads85127 Mouse ato
29	1067	95.7	206	9	ADY74838	Ady74838 Mouse par
30	1062	95.2	206	8	ADI27204	Adi27204 Rat LRP b
31	1058	94.9	206	8	ADI27144	Adi27144 Mouse LRP
32	1055	94.6	206	9	ADY74837	Ady74837 Rat parti
33	992	89.0	183	5	ABB07213	Abb07213 Human clo
34	974	87.4	182	5	AAU82655	Aau82655 Human bre
35	969	86.9	206	8	ADI27151	Adi27151 Chicken L
36	719	64.5	134	6	ABR47438	Abr47438 Breast ca
37	719	64.5	134	8	ADS85025	Ads85025 Human ato
38	510.5	45.8	126	8	ADI27205	Adi27205 Zebrafish
39	482.5	43.3	116	2	AAW44090	Aaw44090 Human sec
40	482.5	43.3	116	2	AAW27654	Aaw27654 Secreted
41	406	36.4	77	8	ADI27214	Adi27214 Mouse LRP
42	355.5	31.9	213	8	ADI27150	Adi27150 Human LRP
43	349.5	31.3	213	3	AAY96429	Aay96429 Human TGF
44	349.5	31.3	213	3	AAY96436	Aay96436 Human TGF
45	349.5	31.3	213	3	AAY96430	Aay96430 Human TGF

ALIGNMENTS

RESULT 1

AAW09408

ID AAW09408 standard; protein; 206 AA.

XX

AC AAW09408;

XX

DT 03-SEP-1997 (first entry)

XX
DE Human small CCN-like growth factor.
XX
KW CCN; human small CNN-like growth factor; SCGF; family; growth regulator;
KW CTGF; connective tissue; cef 10; cry 61; nov; receptor; identify;
KW antagonist; diagnosis; tumour; vascular disease; neovascularisation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .23
FT /label= signal_peptide
FT Protein 24. .206
FT /label= mature_protein
XX
PN WO9639486-A1.
XX
PD 12-DEC-1996.
XX
PF 05-JUN-1995; 95WO-US007092.
XX
PR 05-JUN-1995; 95WO-US007092.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Hastings GA, Adams MD;
XX
DR WPI; 1997-043109/04.
DR N-PSDB; AAT47661.
XX
PT DNA encoding human small CCN-like growth factor polypeptide - useful to
PT diagnose, e.g. tumours.
XX
PS Claim 1; Fig 1; 61pp; English.
XX
CC The present sequence is that of human small CNN-like growth factor
CC (SCGF). SCGF is similar to the CCN family of growth regulators, where CCN
CC stands for CTGF (connective tissue growth factor), cef 10/cry 61, and
CC nov. The SCGF receptor can be used to identify (ant)agonists to SCGF.
CC Disease or a susceptibility to disease related to a mutation of the DNA
CC encoding SCGF (AAT47661) can be diagnosed by determining a mutation in
CC the DNA. Detection of the mutation will allow diagnosis of, e.g. a
CC tumour. An anti-SCGF antibody could be used to diagnose vascular disease
CC or neovascularisation associated with tumour formation
XX
SQ Sequence 206 AA;

Query Match 100.0%; Score 1115; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.3e-109;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPPAIHFYLLPLACILMKSCFLAFKNDATILYSHVVKPVPAPHPSSNSTLNQARNNGGRHF 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MLPPAIHFYLLPLACILMKSCFLAFKNDATILYSHVVKPVPAPHPSSNSTLNQARNNGGRHF 60
Qy 61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
Qy 121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRYKITVVTACKCKRYTRQHNESSHNFE 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

SCORE Search Results Details for Application 10627604 and Search Result us-10-627-604-2.raij.

[Score Home](#)
[Page](#)

Retrieve Application List

SCORE System Overview

SCORE
FAO

Comments /
Suggestions

This page gives you Search Results detail for the Application 10627604 and Search Result us-10-627-604-2.ra1.

start

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

```
Run on:      August 12, 2006, 07:01:10 ; Search time 50 Seconds
              (without alignments)
              360.626 Million cell updates/sec
```

Title: US-10-627-604-2
Perfect score: 1115
Sequence: 1 MLPPAIHFYLLPLACILMKS.....PVOHHRERKRASKSSKHSMS 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :      Issued_Patents_AA:*
1:  /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep:*
2:  /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep:*
3:  /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep:*
4:  /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep:*
5:  /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS_COMB.pep:*
6:  /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep:*
7:  /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
--------	-------	------------------	--------	----	----	-------------

1	1115	100.0	206	1	US-08-468-847B-20	Sequence 20, Appl
2	1109	99.5	206	1	US-08-468-847B-2	Sequence 2, Appli
3	1088	97.6	230	2	US-10-104-047-3602	Sequence 3602, Ap
4	1067	95.7	206	2	US-09-188-930-159	Sequence 159, App
5	1067	95.7	206	2	US-09-188-930-286	Sequence 286, App
6	1067	95.7	206	2	US-09-312-283C-159	Sequence 159, App
7	1067	95.7	206	2	US-09-312-283C-286	Sequence 286, App
8	349.5	31.3	213	2	US-09-449-218D-2	Sequence 2, Appli
9	349.5	31.3	213	2	US-09-449-218D-6	Sequence 6, Appli
10	349.5	31.3	213	2	US-09-449-218D-8	Sequence 8, Appli
11	349.5	31.3	213	2	US-09-668-529A-2	Sequence 2, Appli
12	349.5	31.3	213	2	US-09-668-529A-6	Sequence 6, Appli
13	349.5	31.3	213	2	US-09-668-529A-8	Sequence 8, Appli
14	349.5	31.3	213	2	US-09-668-037A-2	Sequence 2, Appli
15	349.5	31.3	213	2	US-09-668-037A-6	Sequence 6, Appli
16	349.5	31.3	213	2	US-09-668-037A-8	Sequence 8, Appli
17	349.5	31.3	213	2	US-09-668-021-2	Sequence 2, Appli
18	349.5	31.3	213	2	US-09-668-021-6	Sequence 6, Appli
19	349.5	31.3	213	2	US-09-668-021-8	Sequence 8, Appli
20	348	31.2	213	2	US-09-449-218D-10	Sequence 10, Appl
21	348	31.2	213	2	US-09-668-529A-10	Sequence 10, Appl
22	348	31.2	213	2	US-09-668-037A-10	Sequence 10, Appl
23	348	31.2	213	2	US-09-668-021-10	Sequence 10, Appl
24	340	30.5	213	2	US-09-449-218D-14	Sequence 14, Appl
25	340	30.5	213	2	US-09-668-529A-14	Sequence 14, Appl
26	340	30.5	213	2	US-09-668-037A-14	Sequence 14, Appl
27	340	30.5	213	2	US-09-668-021-14	Sequence 14, Appl
28	333	29.9	176	2	US-09-449-218D-16	Sequence 16, Appl
29	333	29.9	176	2	US-09-668-529A-16	Sequence 16, Appl
30	333	29.9	176	2	US-09-668-037A-16	Sequence 16, Appl
31	333	29.9	176	2	US-09-668-021-16	Sequence 16, Appl
32	333	29.9	211	2	US-09-449-218D-12	Sequence 12, Appl
33	333	29.9	211	2	US-09-668-529A-12	Sequence 12, Appl
34	333	29.9	211	2	US-09-668-037A-12	Sequence 12, Appl
35	333	29.9	211	2	US-09-668-021-12	Sequence 12, Appl
36	95.5	8.6	375	1	US-08-468-847B-13	Sequence 13, Appl
37	95.5	8.6	375	2	US-09-495-448A-33	Sequence 33, Appl
38	95	8.5	379	1	US-08-468-847B-11	Sequence 11, Appl
39	95	8.5	379	2	US-09-142-569-2	Sequence 2, Appli
40	95	8.5	379	2	US-09-495-448A-2	Sequence 2, Appli
41	95	8.5	379	2	US-10-099-322-45	Sequence 45, Appl
42	95	8.5	379	2	US-10-044-564-45	Sequence 45, Appl
43	91.5	8.2	735	2	US-09-191-647-9	Sequence 9, Appli
44	91.5	8.2	735	2	US-09-540-245A-9	Sequence 9, Appli
45	91.5	8.2	735	2	US-09-540-153-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-08-468-847B-20

; Sequence 20, Application US/08468847B

; Patent No. 5780263

; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A. and Adams, Mark D.

; TITLE OF INVENTION: Human CCN-Like Growth Factor

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILIAN,

```

;   ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;   STREET:    6 BECKER FARM ROAD
;   CITY:     ROSELAND
;   STATE:    NEW JERSEY
;   COUNTRY:   USA
;   ZIP:      07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  3.5 INCH DISKETTE
;   COMPUTER:   IBM PS/2
;   OPERATING SYSTEM:  MS-DOS
;   SOFTWARE:   WORD PERFECT 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/468,847B
;   FILING DATE:   6 June 1995
;   CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME:   MULLINS, J.G.
;   REGISTRATION NUMBER:  33,073
;   REFERENCE/DOCKET NUMBER:  325800-442
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  201-994-1700
;   TELEFAX:   201-994-1744
;   INFORMATION FOR SEQ ID NO:  20:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:   206 AMINO ACIDS
;   TYPE:     AMINO ACID
;   STRANDEDNESS:
;   TOPOLOGY:  LINEAR
;   MOLECULE TYPE:  PROTEIN
US-08-468-847B-20

```

```

Query Match          100.0%;  Score 1115;  DB 1;  Length 206;
Best Local Similarity 100.0%;  Pred. No. 6.2e-118;
Matches 206;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1  MLPPAIHFYLLPLACILMKSClafKNDaTEILYSHVVKPVPaHPSSNSTLNQARNGGRHF 60
      |||
Db      1  MLPPAIHFYLLPLACILMKSClafKNDaTEILYSHVVKPVPaHPSSNSTLNQARNGGRHF 60

Qy     61  SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
      |||
Db     61  SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120

Qy    121  KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180
      |||
Db    121  KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRTYKITVVTACKCKRYTRQHNESSHNFE 180

Qy    181  SMSPAKPVQHHREKRASKSSKHSMS 206
      |||
Db    181  SMSPAKPVQHHREKRASKSSKHSMS 206

```

RESULT 2

US-08-468-847B-2

; Sequence 2, Application US/08468847B

; Patent No. 5780263

; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A. and Adams, Mark D.

SCORE Search Results Details for Application 10627604 and Search Result us-10-627-604-2.rapbm.

[Score Home](#)
[Page](#)

Retrieve Application List

SCORE System Overview

SCORE
FAO

Comments /
Suggestions

This page gives you Search Results detail for the Application 10627604 and Search Result us-10-627-604-2.rapbm.

start

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

```
Run on:      August 12, 2006, 07:13:05 ; Search time 177 Seconds
              (without alignments)
              539.109 Million cell updates/sec
```

Title: US-10-627-604-2
Perfect score: 1115
Sequence: 1 MLPPAIHFYLLPLACILMKS.....PVOHHRERKRASKSSKHSMS 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :      Published_Applications_AA_Main:*
1:  /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2:  /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
3:  /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
4:  /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
5:  /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
6:  /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%					
Result	Query				
No.	Score	Match	Length	DB	ID
					Description

1	1115	100.0	206	3	US-09-853-625B-20	Sequence 20, Appl
2	1115	100.0	206	4	US-10-223-085-70	Sequence 70, Appl
3	1115	100.0	206	4	US-10-223-084-70	Sequence 70, Appl
4	1115	100.0	206	4	US-10-223-088-70	Sequence 70, Appl
5	1115	100.0	206	4	US-10-223-090-70	Sequence 70, Appl
6	1115	100.0	206	4	US-10-223-087-70	Sequence 70, Appl
7	1115	100.0	206	4	US-10-223-083-70	Sequence 70, Appl
8	1115	100.0	206	4	US-10-223-089-70	Sequence 70, Appl
9	1115	100.0	206	4	US-10-223-081-70	Sequence 70, Appl
10	1115	100.0	206	4	US-10-223-082-70	Sequence 70, Appl
11	1115	100.0	206	4	US-10-305-654-70	Sequence 70, Appl
12	1115	100.0	206	4	US-10-464-368-104	Sequence 104, App
13	1115	100.0	206	4	US-10-081-056-70	Sequence 70, Appl
14	1115	100.0	206	4	US-10-627-604-2	Sequence 2, Appli
15	1115	100.0	206	5	US-10-627-604-2	Sequence 2, Appli
16	1109	99.5	206	3	US-09-853-625B-2	Sequence 2, Appli
17	1109	99.5	206	4	US-10-453-919-119	Sequence 119, App
18	1088	97.6	230	4	US-10-104-047-3602	Sequence 3602, Ap
19	1088	97.6	230	6	US-11-072-512-3602	Sequence 3602, Ap
20	1067	95.7	206	3	US-09-866-050A-159	Sequence 159, App
21	1067	95.7	206	3	US-09-866-050A-286	Sequence 286, App
22	1062	95.2	206	4	US-10-464-368-105	Sequence 105, App
23	1058	94.9	206	4	US-10-464-368-45	Sequence 45, Appl
24	992	89.0	183	3	US-09-867-274-25	Sequence 25, Appl
25	992	89.0	183	4	US-10-679-670A-25	Sequence 25, Appl
26	980	87.9	182	4	US-10-453-919-121	Sequence 121, App
27	969	86.9	206	4	US-10-464-368-52	Sequence 52, Appl
28	719	64.5	134	4	US-10-177-293-108	Sequence 108, App
29	510.5	45.8	126	4	US-10-464-368-106	Sequence 106, App
30	406	36.4	77	4	US-10-464-368-115	Sequence 115, App
31	355.5	31.9	213	4	US-10-464-368-51	Sequence 51, Appl
32	349.5	31.3	213	3	US-09-867-274-5	Sequence 5, Appli
33	349.5	31.3	213	4	US-10-245-752-64	Sequence 64, Appl
34	349.5	31.3	213	4	US-10-245-859-64	Sequence 64, Appl
35	349.5	31.3	213	4	US-10-245-103-64	Sequence 64, Appl
36	349.5	31.3	213	4	US-10-245-107-64	Sequence 64, Appl
37	349.5	31.3	213	4	US-10-245-143-64	Sequence 64, Appl
38	349.5	31.3	213	4	US-10-245-771-64	Sequence 64, Appl
39	349.5	31.3	213	4	US-10-245-851-64	Sequence 64, Appl
40	349.5	31.3	213	4	US-10-245-883-64	Sequence 64, Appl
41	349.5	31.3	213	4	US-10-237-535-64	Sequence 64, Appl
42	349.5	31.3	213	4	US-10-238-183-64	Sequence 64, Appl
43	349.5	31.3	213	4	US-10-238-283-64	Sequence 64, Appl
44	349.5	31.3	213	4	US-10-238-370-64	Sequence 64, Appl
45	349.5	31.3	213	4	US-10-245-055-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1

US-09-853-625B-20

; Sequence 20, Application US/09853625B

; Patent No. US20020049304A1

; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A. and Adams, Mark D.

; TITLE OF INVENTION: Human CCN-Like Growth Factor

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILIAN,
; CECCHI, STEWART & OLSTEIN

```

; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/853,625B
; FILING DATE: 14-May-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/053,587
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-853-625B-20

```

```

Query Match          100.0%; Score 1115; DB 3; Length 206;
Best Local Similarity 100.0%; Pred. No. 2.5e-106;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLPPAIHFYLLPLACILMKSCFLAFKNDATILYSHVVKPVPAPHPSSNSTLNQARNGGRHF 60
|
Db      1 MLPPAIHFYLLPLACILMKSCFLAFKNDATILYSHVVKPVPAPHPSSNSTLNQARNGGRHF 60
|
Qy     61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
|
Db     61 SNTGLDRNTRVQVGCRELRSTKYISDGQCTSI SPLKELVCAGECLPLPVL PNWIGGGYGT 120
|
Qy    121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRYKITVVTACKCKRYTRQHNESSHNFE 180
|
Db    121 KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTRYKITVVTACKCKRYTRQHNESSHNFE 180
|
Qy    181 SMSPAKPVQHHRRERKRASKSSKHSMS 206
|
Db    181 SMSPAKPVQHHRRERKRASKSSKHSMS 206

```

RESULT 2

```

US-10-223-085-70
; Sequence 70, Application US/10223085
; Publication No. US20030100497A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

SCORE Search Results Details for Application 10627604 and Search Result us-10-627-604-2.rapb

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10627604 and Search Result us-10-627-604-2.rapb. [start](#)

[Go Back](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2006, 07:13:40 ; Search time 32 Seconds
(without alignments)
433.321 Million cell updates/sec

Title: US-10-627-604-2
Perfect score: 1115
Sequence: 1 MLPPAIHFYLLPLACILMKS.....PVQHHRRERKRASKSSKHSMS 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 239914 seqs, 67312017 residues

Total number of hits satisfying chosen parameters: 239914

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	118.5	10.6	2233	6 US-10-544-731-9	Sequence 9, Appli
2	102.5	9.2	2202	6 US-10-544-731-8	Sequence 8, Appli

3	102.5	9.2	2227	6	US-10-544-731-7	Sequence 7, Appli
4	89	8.0	984	6	US-10-449-902-56454	Sequence 56454, A
5	84	7.5	250	7	US-11-056-355B-7514	Sequence 7514, Ap
6	82	7.4	5179	7	US-11-105-233-185	Sequence 185, App
7	80	7.2	422	6	US-10-449-902-44214	Sequence 44214, A
8	77	6.9	298	6	US-10-449-902-39708	Sequence 39708, A
9	77	6.9	359	6	US-10-511-937-2932	Sequence 2932, Ap
10	77	6.9	359	6	US-10-511-937-2955	Sequence 2955, Ap
11	77	6.9	500	7	US-11-254-195-6	Sequence 6, Appli
12	76.5	6.9	838	6	US-10-449-902-45252	Sequence 45252, A
13	74.5	6.7	354	7	US-11-056-355B-16795	Sequence 16795, A
14	74.5	6.7	356	7	US-11-056-355B-16794	Sequence 16794, A
15	74.5	6.7	386	7	US-11-056-355B-16793	Sequence 16793, A
16	74.5	6.7	780	6	US-10-449-902-44082	Sequence 44082, A
17	73	6.5	468	7	US-11-293-697-4473	Sequence 4473, Ap
18	72.5	6.5	723	7	US-11-289-102-290	Sequence 290, App
19	71	6.4	215	6	US-10-953-349-31032	Sequence 31032, A
20	71	6.4	224	6	US-10-449-902-37984	Sequence 37984, A
21	71	6.4	224	6	US-10-449-902-49422	Sequence 49422, A
22	71	6.4	627	7	US-11-056-355B-70995	Sequence 70995, A
23	71	6.4	747	7	US-11-056-355B-70994	Sequence 70994, A
24	71	6.4	783	7	US-11-056-355B-70993	Sequence 70993, A
25	70.5	6.3	390	6	US-10-953-349-24421	Sequence 24421, A
26	70.5	6.3	390	7	US-11-056-355B-57508	Sequence 57508, A
27	70.5	6.3	392	6	US-10-953-349-24420	Sequence 24420, A
28	70.5	6.3	392	7	US-11-056-355B-57507	Sequence 57507, A
29	70.5	6.3	590	6	US-10-953-349-9881	Sequence 9881, Ap
30	70.5	6.3	590	7	US-11-056-355B-20740	Sequence 20740, A
31	70.5	6.3	604	6	US-10-953-349-9880	Sequence 9880, Ap
32	70.5	6.3	604	7	US-11-056-355B-20739	Sequence 20739, A
33	70.5	6.3	607	6	US-10-953-349-9879	Sequence 9879, Ap
34	70.5	6.3	607	7	US-11-056-355B-20738	Sequence 20738, A
35	70.5	6.3	847	7	US-11-174-307B-70	Sequence 70, Appl
36	69.5	6.2	204	7	US-11-330-403-8415	Sequence 8415, Ap
37	69.5	6.2	364	7	US-11-056-355B-20081	Sequence 20081, A
38	69.5	6.2	382	7	US-11-056-355B-20080	Sequence 20080, A
39	69.5	6.2	405	7	US-11-056-355B-20079	Sequence 20079, A
40	69.5	6.2	896	7	US-11-056-355B-101044	Sequence 101044,
41	69.5	6.2	896	7	US-11-056-355B-112283	Sequence 112283,
42	69.5	6.2	901	7	US-11-056-355B-101043	Sequence 101043,
43	69.5	6.2	901	7	US-11-056-355B-112282	Sequence 112282,
44	69.5	6.2	902	7	US-11-056-355B-101042	Sequence 101042,
45	69.5	6.2	902	7	US-11-056-355B-112281	Sequence 112281,

ALIGNMENTS

RESULT 1

US-10-544-731-9

; Sequence 9, Application US/10544731

; Publication No. US20060150262A1

; GENERAL INFORMATION:

; APPLICANT: Applied Research Systems ARS Holding N.V.

; TITLE OF INVENTION: NOVEL MUCIN-LIKE POLYPEPTIDES

; FILE REFERENCE: 825-PCT

; CURRENT APPLICATION NUMBER: US/10/544,731

; CURRENT FILING DATE: 2005-08-05

; PRIOR APPLICATION NUMBER: US 60/445,217

; PRIOR FILING DATE: 2003-02-05

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 9
 ; LENGTH: 2233
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-544-731-9

Query Match 10.6%; Score 118.5; DB 6; Length 2233;
 Best Local Similarity 21.6%; Pred. No. 0.0034;
 Matches 51; Conservative 24; Mismatches 94; Indels 67; Gaps 7;

Qy 12 PLACILMKSCS-----AFKNDATEILYSHVVKPVPAPHPSSNSTLN 51
 ||:| |::| : | | | :| | || |
 Db 2005 PLSCSLVRSRIPAPAKGGFTPRWVWGAIIIPAAPADTPSCLGLSTPEPGPMSPSLTSVGA 2064

Qy 52 QARNGGRHFSNTGLDRNTRVQVGCREL-----RSTKYISDGQCTSISP 94
 | | : | : || : | | :| |
 Db 2065 AERLGTGAPLSAQDEARMSKDGCRCFCPPPPPPYQNSTCAVYHRSIIQQQGCSSSEP 2124

Qy 95 LKELVCAGECLPLPVLPNWIGGGYGTKYWSRRSSQEW---CVNDKTRTQRIQLQCQDGS 151
 :: | | | : | : | | : | : | | |
 Db 2125 VRLAYCRGNC-----GDSSSMYSLEGNTVEHRCQCCQELRTSLRNVTLHCTDGS 2173

Qy 152 TRTYKITVVTACKC--KRY---TRQHNE-----SSHNFESMSPAKPVQHH 191
 : | : | | | : | | : | : | | : | |
 Db 2174 SRAFSYTEVEECGCMGRRCPAPGDTQHSEAEPEPSQEAESGSWERGVPVSPMHHH 2229

RESULT 2

US-10-544-731-8
 ; Sequence 8, Application US/10544731
 ; Publication No. US20060150262A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Applied Research Systems ARS Holding N.V.
 ; TITLE OF INVENTION: NOVEL MUCIN-LIKE POLYPEPTIDES
 ; FILE REFERENCE: 825-PCT
 ; CURRENT APPLICATION NUMBER: US/10/544,731
 ; CURRENT FILING DATE: 2005-08-05
 ; PRIOR APPLICATION NUMBER: US 60/445,217
 ; PRIOR FILING DATE: 2003-02-05
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 8
 ; LENGTH: 2202
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-544-731-8

Query Match 9.2%; Score 102.5; DB 6; Length 2202;
 Best Local Similarity 21.0%; Pred. No. 0.14;
 Matches 49; Conservative 24; Mismatches 93; Indels 67; Gaps 7;

Qy 12 PLACILMKSCS-----AFKNDATEILYSHVVKPVPAPHPSSNSTLN 51
 ||:| |::| : | | | :| | || |
 Db 1980 PLSCSLVRSRIPAPAKGGFTPRWVWGAIIIPAAPADTPSCLGLSTPEPGPMSPSLTSVGA 2039

Qy 52 QARNGGRHFSNTGLDRNTRVQVGCREL-----RSTKYISDGQCTSISP 94
 | | : | : || : | | :| |
 Db 2040 AERLGTGAPLSAQDEARMSKDGCRCFCPPPPPPYQNSTCAVYHRSIIQQQGCSSSEP 2099

Qy 95 LKELVCAGECLPLPVLPNWIGGGYGTKYWSRRSSQEW---CVNDKTRTQRIQLQCQDGS 151

SCORE Search Results Details for Application 10627604 and Search Result us-10-627-604-2.rapm.

Score Home	Retrieve Application	SCORE System	SCORE	Comments /
Page	List	Overview	FAQ	Suggestions

This page gives you Search Results detail for the Application 10627604 and Search Result us-10-627-604-2.rapm.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2006, 07:01:50 ; Search time 606 Seconds
(without alignments)
519.064 Million cell updates/sec

Title: US-10-627-604-2
Perfect score: 1115
Sequence: 1 MLPPAIHFYLLPLACILMKS.....PVQHHRRERKRASKSSKHSMS 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCTUS_COMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066_COMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073_COMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/paa/US074_COMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075_COMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076_COMB.pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077_COMB.pep:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078_COMB.pep:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079_COMB.pep:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080_COMB.pep:*
- 11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081_COMB.pep:*
- 12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082_COMB.pep:*
- 13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083_COMB.pep:*
- 14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084_COMB.pep:*
- 15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085_COMB.pep:*
- 16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086_COMB.pep:*

```

17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087_COMB.pep:*
18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088_COMB.pep:*
19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089_COMB.pep:*
20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090_COMB.pep:*
21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091_COMB.pep:*
22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092_COMB.pep:*
23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093_COMB.pep:*
24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094_COMB.pep:*
25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095_COMB.pep:*
26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096_COMB.pep:*
27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097_COMB.pep:*
28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098_COMB.pep:*
29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099_COMB.pep:*
30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100_COMB.pep:*
31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101_COMB.pep:*
32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102_COMB.pep:*
33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103_COMB.pep:*
34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104_COMB.pep:*
35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105_COMB.pep:*
36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106_COMB.pep:*
37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107_COMB.pep:*
38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108_COMB.pep:*
39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109_COMB.pep:*
40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110_COMB.pep:*
41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111_COMB.pep:*
42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112_COMB.pep:*
43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113_COMB.pep:*
44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114_COMB.pep:*
45: /EMC_Celerra_SIDS3/ptodata/2/paa/US600_COMB.pep:*
46: /EMC_Celerra_SIDS3/ptodata/2/paa/US601_COMB.pep:*
47: /EMC_Celerra_SIDS3/ptodata/2/paa/US602_COMB.pep:*
48: /EMC_Celerra_SIDS3/ptodata/2/paa/US603_COMB.pep:*
49: /EMC_Celerra_SIDS3/ptodata/2/paa/US604_COMB.pep:*
50: /EMC_Celerra_SIDS3/ptodata/2/paa/US605_COMB.pep:*
51: /EMC_Celerra_SIDS3/ptodata/2/paa/US606_COMB.pep:*
52: /EMC_Celerra_SIDS3/ptodata/2/paa/US607_COMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1115	100.0	206	1	PCT-US03-17409-213	Sequence 213, App
2	1115	100.0	206	1	PCT-US04-24424-2058	Sequence 2058, Ap
3	1115	100.0	206	1	PCT-US98-13530-12	Sequence 12, Appl
4	1115	100.0	206	1	PCT-US99-31005-17	Sequence 17, Appl
5	1115	100.0	206	14	US-08-468-847-2	Sequence 2, Appli
6	1115	100.0	206	18	US-08-885-610-12	Sequence 12, Appl
7	1115	100.0	206	20	US-09-053-587-20	Sequence 20, Appl
8	1115	100.0	206	22	US-09-220-876-17	Sequence 17, Appl
9	1115	100.0	206	28	US-09-850-726-17	Sequence 17, Appl
10	1115	100.0	206	28	US-09-853-625-20	Sequence 20, Appl
11	1115	100.0	206	28	US-09-853-625B-20	Sequence 20, Appl
12	1115	100.0	206	30	US-10-081-056-70	Sequence 70, Appl
13	1115	100.0	206	31	US-10-170-205E-29147	Sequence 29147, A
14	1115	100.0	206	32	US-10-223-081-70	Sequence 70, Appl
15	1115	100.0	206	32	US-10-223-082-70	Sequence 70, Appl

16	1115	100.0	206	32	US-10-223-083-70	Sequence 70, Appl
17	1115	100.0	206	32	US-10-223-084-70	Sequence 70, Appl
18	1115	100.0	206	32	US-10-223-085-70	Sequence 70, Appl
19	1115	100.0	206	32	US-10-223-087-70	Sequence 70, Appl
20	1115	100.0	206	32	US-10-223-088-70	Sequence 70, Appl
21	1115	100.0	206	32	US-10-223-089-70	Sequence 70, Appl
22	1115	100.0	206	32	US-10-223-090-70	Sequence 70, Appl
23	1115	100.0	206	33	US-10-305-654-70	Sequence 70, Appl
24	1115	100.0	206	34	US-10-405-027-3237	Sequence 3237, Ap
25	1115	100.0	206	34	US-10-405-027-4573	Sequence 4573, Ap
26	1115	100.0	206	34	US-10-464-368-104	Sequence 104, App
27	1115	100.0	206	35	US-10-567-867-2058	Sequence 2058, Ap
28	1115	100.0	206	36	US-10-627-604-2	Sequence 2, Appli
29	1115	100.0	206	37	US-10-743-643-2071	Sequence 2071, Ap
30	1115	100.0	206	49	US-60-490-890-2058	Sequence 2058, Ap
31	1115	100.0	223	27	US-09-758-444-401	Sequence 401, App
32	1115	100.0	223	32	US-10-212-160-401	Sequence 401, App
33	1109	99.5	206	20	US-09-053-587-2	Sequence 2, Appli
34	1109	99.5	206	27	US-09-778-381-119	Sequence 119, App
35	1109	99.5	206	28	US-09-853-625-2	Sequence 2, Appli
36	1109	99.5	206	28	US-09-853-625B-2	Sequence 2, Appli
37	1109	99.5	206	34	US-10-453-919-119	Sequence 119, App
38	1108	99.4	206	28	US-09-834-366-19284	Sequence 19284, A
39	1108	99.4	206	46	US-60-197-873-19284	Sequence 19284, A
40	1100.5	98.7	205	28	US-09-853-625C-2	Sequence 2, Appli
41	1088	97.6	230	40	US-11-072-512-3602	Sequence 3602, Ap
42	1067	95.7	206	20	US-09-069-726-159	Sequence 159, App
43	1067	95.7	206	28	US-09-866-050A-159	Sequence 159, App
44	1067	95.7	206	28	US-09-866-050A-286	Sequence 286, App
45	1062	95.2	206	34	US-10-464-368-105	Sequence 105, App

ALIGNMENTS

RESULT 1

PCT-US03-17409-213

; Sequence 213, Application PC/TUS0317409

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION

; APPLICANT: RICKERT, Paula K.

; APPLICANT: KRASNOW, Randi

; TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR LUNG CANCER

; FILE REFERENCE: PA-0051 PCT

; CURRENT APPLICATION NUMBER: PCT/US03/17409

; CURRENT FILING DATE: 2003-06-02

; PRIOR APPLICATION NUMBER: US 60/386,005

; PRIOR FILING DATE: 2002-06-04

; NUMBER OF SEQ ID NOS: 296

; SOFTWARE: PERL Program

; SEQ ID NO 213

; LENGTH: 206

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No: 1420377CD1

PCT-US03-17409-213

Query Match 100.0%; Score 1115; DB 1; Length 206;

Best Local Similarity 100.0%; Pred. No. 1.3e-118;

SCORE Search Results Details for Application 10627604 and Search Result us-10-627-604- 2.rapn.

Score Home	Retrieve Application	SCORE System	SCORE	Comments /
Page	List	Overview	FAQ	Suggestions

This page gives you Search Results detail for the Application 10627604 and Search Result us-10-627-604-2.rapn.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2006, 07:02:50 ; Search time 35 Seconds
(without alignments)
393.959 Million cell updates/sec

Title: US-10-627-604-2
Perfect score: 1115
Sequence: 1 MLPPAIHFYLLPLACILMKS.....PVQHHRRERKRASKSSKHSMS 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 298239 seqs, 66934740 residues

Total number of hits satisfying chosen parameters: 298239

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCT_NEW_COMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/paa/US06_NEW_COMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/paa/US07_NEW_COMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/paa/US08_NEW_COMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/paa/US09_NEW_COMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/paa/US10_NEW_COMB.pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/paa/US11_NEW_COMB.pep:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query
--------	---	-------

No.	Score	Match	Length	DB	ID	Description
1	341	30.6	190	7	US-11-411-003-1	Sequence 1, Appli
2	340	30.5	213	7	US-11-411-003-98	Sequence 98, Appl
3	95.5	8.6	375	7	US-11-455-506-20	Sequence 20, Appl
4	95	8.5	379	1	PCT-US06-11960-555	Sequence 555, App
5	95	8.5	379	6	US-10-182-432A-2	Sequence 2, Appli
6	95	8.5	379	7	US-11-455-506-2	Sequence 2, Appli
7	90	8.1	381	1	PCT-US06-05812-1	Sequence 1, Appli
8	90	8.1	381	6	US-10-182-432A-4	Sequence 4, Appli
9	90	8.1	381	7	US-11-455-506-4	Sequence 4, Appli
10	82	7.4	1048	7	US-11-475-062-6109	Sequence 6109, Ap
11	82	7.4	1490	7	US-11-475-062-6110	Sequence 6110, Ap
12	82	7.4	2839	7	US-11-475-062-6111	Sequence 6111, Ap
13	82	7.4	5179	7	US-11-318-418-185	Sequence 185, App
14	82	7.4	5179	7	US-11-318-418-185	Sequence 185, App
15	82	7.4	5179	7	US-11-437-729-1766	Sequence 1766, Ap
16	82	7.4	5179	7	US-11-475-062-2277	Sequence 2277, Ap
17	81.5	7.3	393	1	PCT-US06-18535-32505	Sequence 32505, A
18	81.5	7.3	393	7	US-11-431-855-32505	Sequence 32505, A
19	81.5	7.3	393	8	US-60-815-535-199	Sequence 199, App
20	79	7.1	353	1	PCT-US06-18535-5233	Sequence 5233, Ap
21	79	7.1	353	7	US-11-431-855-5233	Sequence 5233, Ap
22	77.5	7.0	342	1	PCT-US06-18535-22654	Sequence 22654, A
23	77.5	7.0	342	7	US-11-431-855-22654	Sequence 22654, A
24	77	6.9	426	6	US-10-534-579-20	Sequence 20, Appl
25	77	6.9	500	7	US-11-490-374-377	Sequence 377, App
26	77	6.9	500	7	US-11-490-374-378	Sequence 378, App
27	77	6.9	500	7	US-11-490-374-379	Sequence 379, App
28	76	6.8	604	7	US-11-345-403-12	Sequence 12, Appl
29	75.5	6.8	1150	1	PCT-US06-18535-24537	Sequence 24537, A
30	75.5	6.8	1150	7	US-11-431-855-24537	Sequence 24537, A
31	75	6.7	445	7	US-11-490-374-376	Sequence 376, App
32	74.5	6.7	339	7	US-11-488-375-32	Sequence 32, Appl
33	74.5	6.7	339	7	US-11-488-375-36	Sequence 36, Appl
34	74.5	6.7	354	7	US-11-488-375-37	Sequence 37, Appl
35	74.5	6.7	372	6	US-10-533-520-4748	Sequence 4748, Ap
36	74.5	6.7	372	7	US-11-488-375-33	Sequence 33, Appl
37	74.5	6.7	402	7	US-11-475-062-5674	Sequence 5674, Ap
38	74.5	6.7	428	7	US-11-475-062-5676	Sequence 5676, Ap
39	74.5	6.7	428	7	US-11-475-062-5678	Sequence 5678, Ap
40	74.5	6.7	428	7	US-11-475-062-5683	Sequence 5683, Ap
41	74.5	6.7	1686	1	PCT-US06-04985-1	Sequence 1, Appli
42	73	6.5	555	6	US-10-533-520-3634	Sequence 3634, Ap
43	72.5	6.5	698	7	US-11-336-706-15	Sequence 15, Appl
44	72.5	6.5	1038	6	US-10-670-914-91	Sequence 91, Appl
45	72	6.5	467	6	US-10-311-974A-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-11-411-003-1

; Sequence 1, Application US/11411003

; GENERAL INFORMATION:

; APPLICANT: Paszty, Christopher

; APPLICANT: Robinson, Martyn Kim

; APPLICANT: Graham, Kevin

; APPLICANT: Henry, Alistair James

; APPLICANT: Hoffmann, Kelly Sue

Query Match 30.6%; Score 341; DB 7; Length 190;
Best Local Similarity 42.7%; Pred. No. 7.3e-27;
Matches 79; Conservative 30; Mismatches 62; Indels 14; Gaps 8;

```

RESULT 2
US-11-411-003-98
; Sequence 98, Application US/11411003
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Robinson, Martyn Kim
; APPLICANT: Graham, Kevin
; APPLICANT: Henry, Alistair James
; APPLICANT: Hoffmann, Kelly Sue
; APPLICANT: Latham, John
; APPLICANT: Lawson, Alastair
; APPLICANT: Lu, Hsieng Sen
; APPLICANT: Popplewell, Andy

```